PENGGUNAAN METODE BOOTSTRAP UNTUK PEMODELAN FUNGSI RELIABILITAS DISTRIBUSI GUMBEL TERSENSOR TYPE III

Ardi Kurniawan[1)]

Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga
Kampus C Unair Jalan Mulyorejo, Surabaya
[1)]ardik2008@gmail.com

Abstract—**Pada penelitian penentuan lama tahan hidup suatu benda atau individu, kadang kala tidak dapat diamati lama tahan hidup seluruh sampel penelitian yang digunakan. Hal tersebut bisa diakibatkan oleh keterbatasan waktu, ketidakcukupan biaya, dan kurangnya tenaga di dalam penelitian yang dilakukan. Penyensoran sampel penelitian merupakan salah satu cara yang dapat dipakai untuk mengatasi sulitnya mengamati lama tahan hidup seluruh sampel dalam periode waktu penelitian yang telah ditentukan. Pada sampel tersensor type III, obyek penelitian masuk dalam pengujian pada waktu yang tidak bersamaan. Obyek penelitian yang mati atau gagal sebelum waktu pengamatan berakhir dikatakan tidak tersensor, sedangkan obyek penelitian yang melebihi batas waktu penelitian yang diberikan atau mengundurkan diri sebelum waktu penelitian berakhir disebut tersensor. Distribusi Gumbel adalah distribusi yang banyak digunakan dalam pembahasan Analisis Data Uji Hidup. Distribusi ini mempunyai dua parameter, yaitu parameter Bentuk dan Lokasi. Berdasarkan distribusi Gumbel untuk kasus tersensor type III dicari nilai penduga parameter distribusi Gumbel tersebut melalui proses resampling dengan menggunakan metode Bootstrap. Selain itu ditentukan pula bentuk fungsi reliabilitas dari distribusi Gumbel. Data yang digunakan dalam penelitian ini adalah data penderita kanker yang dirawat di RSUP Dr. Karyadi Semarang. Selanjutnya dihitung peluang hidup penderita kanker tersebut memakai Fungsi Reliabilitas yang diperoleh.**

Kata Kunci : **Distribusi Gumbel, Sampel Tersensor Type III, Metode Bootstrap, Fungsi Reliabilitas**

I. PENDAHULUAN

Distribusi Gumbel merupakan suatu distribusi yang ditemukan oleh seorang ahli statistika E.J. Gumbel (1958). Distribusi Gumbel ini mempunyai dua parameter, yaitu parameter lokasi dan parameter skala. Distribusi ini biasanya dikenal pula dengan nama Distribusi Nilai Ekstrim. Distribusi Gumbel erat hubungannya dengan distribusi Weibull. Eksponen dari distribusi Gumbel merupakan distribusi Weibull (Lawless, 2003).

Pada penelitian Uji Hidup, data waktu hidup yang diperoleh dapat berupa data tersensor dan data tidak tersensor. Data tersensor terjadi jika lama tahan hidup seluruh sampel yang digunakan dalam penelitian tidak diketahui secara pasti. Hal tersebut bisa diakibatkan oleh keterbatasan waktu, ketidakcukupan biaya, dan kurangnya tenaga di dalam penelitian yang dilakukan. Salah satu jenis penyensoran yang sering digunakan dalam analisis lama tahan hidup adalah sampel tersensor type III. Suatu penyensoran dikatakan tersensor type III apabila sampel masuk ke dalam percobaan dalam waktu yang tidak sama.

Kadang kala permasalahan yang sering timbul pada penelitian data Uji Hidup adalah kecilnya jumlah sampel yang diperoleh. Salah penyelesaian untuk mengatasi permasalahan ini adalah digunakannya Metode Bootstrap. Metode ini diperkenalkan oleh seorang ahli statistika yang bernama Efron pada tahun 1970. Metode Bootstrap merupakan metode berbasis resampling data sampel dengan syarat adanya pengembalian.

Penelitian yang terkait dalam analisis data uji hidup yang menggunakan metode Bootstrap diantaranya dilakukan oleh Rahayu dkk (2012) dan Solih dkk (2017). Rahayu dkk (2012) menggunakan metode Bootstrap untuk pemodelan Regresi Cox Proportional Hazards pada Ketahanan Hidup Pasien Diabetis Mellitus. Sedangkan Solih dkk (2017) meneliti tentang penggunaan Metode Bootstrap untuk pengestimasian Interval Kepercayaan parameter model proses Geometrik Weibull pada data tersensor type II.

Berdasarkan uraian di atas penulis ingin mengembangkan lebih jauh penelitian yang terkait dengan analisis data uji hidup dan metode Bootstrap. Distribusi data yang dipakai adalah distribusi Gumbel dan metode penyensoran yang digunakan adalah metode penyensoran type III. Selanjutnya diestimasi bentuk fungsi Reliabilitasnya dan diterapkan pada Data Pasien Kanker Paru yang dirawat di RSUP Karyadi Semarang tahun 2015.

II. TINJAUAN PUSTAKA

2.1. Sampel Tersensor III

Pengamatan waktu tahan hidup sampel terhadap suatu perlakuan sering menjadi fokus dalam suatu penelitian. Kadang kala terdapat sampel yang waktu tahan hidupnya diketahui sebelum pengamatan berakhir, atau kemungkinan kedua adalah sampel keluar sebelum pengamatan berakhir, atau kemungkinan ketiga adalah sampel tetap hidup hingga batas berakhirnya penelitian. Sampel tersensor adalah pengamatan waktu tahan hidup sampel percobaan yang tidak dapat dilakukan sampai selesai akan diperoleh sampel tersensor. Penyebab digunakannya sampel tersensor diantaranya adalah keterbatasan waktu, ketidakcukupan biaya, dan kurangnya tenaga di dalam penelitian yang dilakukan.

Terdapat beberapa jenis sampel yang sering digunakan dalam eksperimen data uji hidup, salah satu diantaranya adalah sampel tersensor type III. Jenis sampel ini diperoleh jika individu atau objek masuk ke dalam penelitian pada waktu yang berlainan selama periode waktu yang telah ditetapkan. Fungsi likelihood pada data tersensor tipe III dari data pengamatan (t_i, δ_i) , $i = 1, 2, \ldots, n$ yaitu

$$L(\theta) = \prod_{i=1}^{n} f(t_i)^{\delta_i} \; S(t_i)^{1-\delta_i}$$

dengan :
δ_i adalah indikator penyensoran, δ_i bernilai 1 jika data tidak tersensor dan δ_i bernilai 0 jika data tersensor.
t_i diperoleh dari $\min(T_i, C_i), i = 1, 2, \ldots, n$.
T_i adalah waktu hidup individu dan C_i adalah waktu penyensoran.

2.2. Distribusi Gumbel

Distribusi Gumbel atau distribusi Nilai Ekstrim merupakan suatu distribusi yang sering digunakan dalam masalah Analisis Data Uji Hidup. Fungsi kepadatan probabilitas (fkp) distribusi Gumbel diberikan dengan :

$$f(y) = (1/b) \exp[(y-u)/b - \exp\{(y-u)/b)\}] \qquad (1)$$

atau

$$f(y) = (1/b) \exp((y-u)/b) \exp(- \exp((y-u)/b) \qquad (2)$$

dengan : $-\infty < y < \infty$, u adalah parameter lokasi, dan b adalah parameter skala.

Bentuk fungsi distribusi kumulatif F(y) dari distribusi Gumbel tersebut tanpa adanya penyensoran adalah :

$$F(y) = 1 - \exp\{- \exp[(y-u)/b]\} \qquad (3)$$

2.3. Metode Bootstrap

Misalkan dalam suatu penelitian dipunyai sampel yang berukuran *n* data amatan, T = (t_1, t_2,…, t_n) dan dari sampel ini dihitung estimator $\hat{\theta} = f(T)$. Metode *Bootstrap* merupakan prosedur pengambilan sampel baru secara berulang sebanyak *B* sampel dari data asal berukuran *n*. Sebuah sampel baru diperoleh melalui pengambilan titik sampel dari data asal secara satu persatu dengan pengembalian.

Bentuk sampel yang diperoleh dari metode Bootstrap dapat digambarkan sebagai berikut :

$$\begin{array}{lllllll}
t_1^{*(1)}, & t_2^{*(1)}, & t_3^{*(1)}, & . & . & . & t_n^{*(1)} \\
t_1^{*(2)}, & t_2^{*(2)}, & t_3^{*(2)}, & . & . & . & t_n^{*(2)} \\
t_1^{*(3)}, & t_2^{*(3)}, & t_3^{*(3)}, & . & . & . & t_n^{*(3)} \\
. & & & & & & \\
. & & & & & & \\
. & & & & & & \\
t_1^{*(B)}, & t_2^{*(B)}, & t_3^{*(B)}, & . & . & . & t_n^{*(B)}
\end{array}$$

Dengan menerapkan metode Bootstrap maka dari sampel baru dan akan diperoleh estimator-estimator $\hat{\theta}^{*j}$ dengan j = 1, 2, … , B. Selanjutnya dari estimator-estimator tersebut dirata-rata untuk menentukan estimator Bootstrap dari distribusi yang diiinginkan.

III. METODE PENELITIAN

Langkah-langkah penelitian yang dilakukan adalah :

1. Mengestimasi parameter titik distribusi Gumbel dengan cara :
   a. Menentukan Estimasi Titik Parameter Distribusi Weibull pada Data Tahan Hidup Tersensor Tipe III dengan Metode *Maximum Likelihood Estimator* (MLE) dan Metode Bootstrap.
   b. Mentransformasi hasil (a) ke dalam bentuk parameter distribusi Gumbel.
2. Menentukan bentuk fungsi reliabilitas distribusi Gumbel pada Data Tahan Hidup Tersensor Type III.
3. Menggunakan hasil fungsi reliabilitas distribusi Gumbel pada lagkah (3) untuk memprediksi probabilitas waktu hidup Pasien Kanker Paru yang dirawat di RSUP Karyadi Semarang.

IV. HASIL DAN PEMBAHASAN

4.1. Estimasi Titik TersensorTipe III

Andaikan variabel random waktu Y berdistribusi Gumbel yang dinyatakan sebagai :

$$f(y) = (1/b) \exp[(y - u)/b] \exp[- \exp\{(y-u)/b)\}],$$

dengan $y > 0$

Dilakukan tranformasi variabel dengan mengambil :

$$T = \mathrm{Exp}(Y) \; ; \; u = \ln(\alpha) \; ; \; \text{dan } b = 1/\beta.$$

Berdasarkan transformasi variabel tersebut diperoleh jacobian yang berbentuk :

$$J = \frac{dy}{dt} = \frac{1}{t}$$

serta fungsi hasil transformasi yang berbentuk :

$$g(t) = f(\ln(t)) \mid J \mid$$

atau

$$g(t) = \beta \exp[(\ln(t) - \ln(\alpha))\beta] \times \exp[-\exp\{-(\ln(t) - \ln(\alpha))\beta\}]\ (1/t)$$

atau

$$g(t) = \frac{\beta}{t}\left(\frac{t}{\alpha}\right)^{\beta} exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right)$$

atau

$$g(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right) \qquad (4)$$

dengan $t > 0$.

Bentuk pdf dari persamaan (4) sering disebut dengan disribusi Weibull. Bentuk CDF G(t) dan Fungsi Reliabilitas S(t) persamaan (4) masing-masing adalah :

$$G(t) = 1 - exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right)$$

dan

$$S(t) = exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right) \qquad (5)$$

Misalkan $T_1, T_2, \dots, T_n$ adalah suatu peubah acak tahan hidup yang berdistribusi identik seperti pada persamaan (6) dan saling bebas. Fungsi likelihood tersensor tipe III dari $T_1, T_2, \dots, T_n$ yaitu :

$$L(\alpha,\beta) = \prod_{i=1}^{n}\left(g(t_i)^{\delta_i} S(t_i)^{1-\delta_i}\right)$$

atau

$$L(\alpha,\beta) = \prod_{i=1}^{n}\left(\left(\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{\delta_i} \left(exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{1-\delta_i}\right) \qquad (6)$$

Bentuk fungsi ln-likelihood persamaan (6) yaitu :

$$\ln(L(\alpha,\beta)) = \sum_{i=1}^{n} ln\left(\left(\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{\delta_i} \left(exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{1-\delta_i}\right) \qquad (7)$$

Untuk memudahkan penulisan, selanjutnya bentuk $\ln(L(\alpha, \beta))$ akan dituliskan dengan lnL. Dengan demikian diperoleh :

$$lnL = \sum_{i=1}^{n} ln\left(\left(\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{\delta_i} \left(exp\left(-\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right)^{1-\delta_i}\right)$$

atau

$$lnL = \sum_{i=1}^{n} \delta_i ln\,\beta + \sum_{i=1}^{n} \beta\delta_i\, ln\, t_i - \sum_{i=1}^{n} \delta_i\, ln\, t_i - \sum_{i=1}^{n} \beta\delta_i \ln\alpha - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} \qquad (8)$$

Turunan lnL terhadap parameter α persamaan (8) dan disamadengankan nol didapatkan :

$$\frac{\partial\, lnL}{\partial\alpha} = \frac{\sum_{i=1}^{n}\delta_i\beta}{\alpha} - \sum_{i=1}^{n} t_i^{\beta}\left(-\beta\alpha^{-\beta-1}\right) = 0$$

Dengan demikian

$$\alpha^{\beta} = \frac{\sum_{i=1}^{n} t_i^{\beta}}{\sum_{i=1}^{n}\delta_i}$$

atau

$$\hat{\alpha} = \left[\frac{\sum_{i=1}^{n} t_i^{\beta}}{\sum_{i=1}^{n}\delta_i}\right]^{\frac{1}{\beta}} \qquad (9)$$

Pada sisi lain, turunan lnL terhadap parameter β persamaan (9) dan disamadengankan nol didapatkan :

$$\frac{\partial\, lnL}{\partial\beta} = \frac{\sum_{i=1}^{n}\delta_i}{\beta} + \sum_{i=1}^{n}\delta_i \ln t_i - \sum_{i=1}^{n}\delta_i \ln\alpha - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} ln\left(\frac{t_i}{\alpha}\right) = 0$$

atau

$$\frac{\sum_{i=1}^{n}\delta_i}{\beta} + \sum_{i=1}^{n}\delta_i \ln t_i - \sum_{i=1}^{n}\delta_i \ln\alpha - \frac{1}{\alpha^{\beta}}\sum_{i=1}^{n} t_i^{\beta} ln(t_i) + \frac{1}{\alpha^{\beta}}\sum_{i=1}^{n} t_i^{\beta} ln(\alpha) = 0 \qquad (10)$$

Mensubstitusikan nilai dugaan α dari persamaan (9) ke dalam persamaan (10) didapatkan :

$$\left(\frac{\sum_{i=1}^{n}\delta_i}{\sum_{i=1}^{n}y_i^{\beta}}\right)\sum_{i=1}^{n}y_i^{\beta}\ln y_i-\frac{\sum_{i=1}^{n}\delta_i}{\beta}$$

$$-\sum_{i=1}^{n}\delta_i\,\ln t_i=0 \qquad (11)$$

Persamaan (11) tersebut merupakan persamaan implisit terhadap parameter β, sehingga nilai $\hat{\beta}$ tidak dapat ditentukan secara langsung. Nilai $\hat{\beta}$ memerlukan pendekatan numerik untuk memperolehnya.
Berdasarkan hasil $\hat{\alpha}$ dan $\hat{\beta}$, maka nilai dugaan parameter u dan b dari distribusi Gumbel masing-masing yaitu : u = ln($\hat{\alpha}$) ; dan b = 1/ $\hat{\beta}$.

4.2. Penerapan Metode Bootstrap

Penerapan dari metode ini digunakan Data Pasien Kanker Paru yang dirawat di RSUP Karyadi Semarang tahun 2015 (Utami, 2015). Data yang digunakan menyatakan waktu kegagalan/kematian yang diambil secara tersensor type III dari 50 pasien penderita kanker paru. Uji Goodness of Fit dengan menggunakan software Easy Fit pada tingkat signifikansi 5% diperoleh, bahwa data berdistribusi Weibull.

Penggunaan metode Bootstrap pada data dilakukan dengan mengambil B = 50 dan didapatkan penduga estimator α serta β masing-masing adalah $\hat{\alpha}=94{,}78$ dan $\hat{\beta}=1{,}64$. Berdasarkan hasil $\hat{\alpha}$ dan $\hat{\beta}$ tersebut, maka diperoleh nilai penduga parameter u dan b dari distribusi Gumbel tersensor type III masing-masing adalah $\hat{u}=\ln(\hat{\alpha})=4{,}55$ dan $\hat{b}=\frac{1}{\hat{\beta}}=0{,}61$.
Dengan demikian bentuk fungsi survivor distribusi Gumbel tersensor type III dari pasien Pasien Kanker Paru RSUP Karyadi Semarang yaitu :

S(y) = exp{- exp[(y-u)/b]}

atau

S(y) = exp{- exp[(y – 4,55)/0,61]}

Sebagai suatu contoh, misalkan ada penderita kanker paru yang telah dirawat selama 100 hari di RSUP Karyadi Semarang, maka probabilitas penderita kanker paru tersebut dapat bertahan hidup lebih dari 100 hari yaitu :

S(y) = exp{- exp[(y – 4,55)/0,61]}
= exp{- exp[(ln(100) – 4,55)/0,61]}
= 0,33

Sama seperti kasus di atas, probabilitas penderita kanker paru mampu bertahan hidup lebih dari 150 hari dan 200 hari masing-masing adalah 0,12 dan 0,033.

V. KESIMPULAN

a. Penduga parameter lokasi dan paramater skala distribusi Gumbel masing-masing adalah $\hat{u}=\left(\left[\frac{\sum_{i=1}^{n}t_i^{\beta}}{\sum_{i=1}^{n}\delta_i}\right]^{\frac{1}{\beta}}\right)$ dan $\hat{b}=\frac{1}{\hat{\beta}}$ dengan $\hat{\beta}$ merupakan penyelesaiaan numerik dari :
$$\left(\frac{\sum_{i=1}^{n}\delta_i}{\sum_{i=1}^{n}y_i^{\beta}}\right)\sum_{i=1}^{n}y_i^{\beta}\ln y_i-\frac{\sum_{i=1}^{n}\delta_i}{\beta}$$
$$-\sum_{i=1}^{n}\delta_i\,\ln t_i=0$$

b. Bentuk fungsi survivor distribusi Gumbel tersensor type III dari pasien Pasien Kanker Paru RSUP Karyadi Semarang yaitu

S(y) = exp{- exp[(y – 4,55)/0,61]}.

c. Semakin lamanya perawatan di rumah sakit harus dijalani oleh penderita kanker, menunjukkan semakin parahnya penyakit kanker paru yang diderita. Hal tersebut ditunjukkan dengan semakin rendahnya probabilitas bertahan hidup si penderita.

VI. PUSTAKA

Lawless, J.F., 2003, *Statistical Models and Methods for Life Time Data*, Jhon Wiley and Sons, New York.

Rahayu N., Setiawan A., dan Mahatma T., 2012, Penggunaan Metode Bootstrap dalam Regresi Cox Proportional Hazards pada Ketahanan Hidup Pasien Diabetis Mellitus, *Prosiding Sendikmad*, Universitas Kristen Satya Wacana.

Solih AA., Cahyandari R., dan Tarkinih, 2017, Estimasi Interval Kepercayaan Parameter Model Proses Geometrik Weibull pada Analisis Uji Hidup untuk Data Tersensor Type II, *Jurnal UIN Sunan Gunung Djati*, Vol X, no 1.

Utami, D.T., 2015, Analisis Data Uji Hidup Pasien Kanker Paru Di RSUP DR. Kariadi Semarang, *Skripsi*, FMIPA, UniversitasNegeri Semarang.